Banco Central de Chile
Documentos de Trabajo

Central Bank of Chile
Working Papers

N° 588

Julio 2010

NOWCASTING WITH GOOGLE TRENDS IN AN EMERGING MARKET

Yan Carrière-Swallow   Felipe Labbé

BANCO CENTRAL DE CHILE

CENTRAL BANK OF CHILE

Documentos de Trabajo del Banco Central de Chile
Working Papers of the Central Bank of Chile
Agustinas 1180, Santiago, Chile
Teléfono: (56-2) 3882475; Fax: (56-2) 3882231

Documento de Trabajo
N° 588

Working Paper
N° 588

NOWCASTING WITH GOOGLE TRENDS IN AN EMERGING MARKET

Yan Carrière-Swallow
Gerencia de Investigación Económica
Banco Central de Chile

Felipe Labbé
Gerencia de Investigación Económica
Banco Central de Chile

Resumen

Gran parte de las variables macro agregadas son reveladas con cierto rezago dificultando así la evaluación del estado actual de la economía. Este trabajo explora si la información sobre búsquedas en Internet puede reflejar el comportamiento de una variable macro en tiempo real en un mercado emergente. Utilizando datos de búsquedas en Google construimos un índice que refleja la intención de comprar un auto, y testeamos si este índice logra mejorar el ajuste y la eficiencia de modelos de *nowcasting* para ventas de autos en Chile. Además, evaluamos si dicho índice proporciona información sobre la predicción de cambios de tendencias. Pese a una relativamente baja penetración de Internet, nuestros resultados indican que los modelos que incorporan nuestro índice tienen varias ventajas: generan predicciones más acertadas en términos de ajuste y eficiencia, dentro y fuera de muestra. Adicionalmente, se muestra cierta ganancia sensibilizando la oportunidad de la entrega de la información.

Abstract

Most economic variables are released with a lag, making it difficult for policy-makers to make an accurate assessment of current conditions. This paper explores whether observing Internet browsing habits can inform practitioners about real-time aggregate consumer behavior in an emerging market. Using data on Google search queries, we introduce a simple index of interest in automobile purchases in Chile and test whether it improves the fit and efficiency of nowcasting models for automobile sales. We also examine to what extent our index helps us identify turning points in sales data. Despite relatively low rates of Internet usage among the population, we find that models incorporating our Google Trends Automotive Index outperform benchmark specifications in both in-sample and out-of-sample nowcasts while providing substantial gains in information delivery times.

We appreciate the helpful comments of Luis Felipe Céspedes, Miguel Ricaurte, Pablo Pincheira, and seminar participants at the Central Bank of Chile. Remaining errors are entirely our own. Emails: ycarriere@bcentral.cl; flabbe@bcentral.cl

1 Introduction

The delay in data releases of key macroeconomic variables presents a limitation for decision-makers by restricting their ability to accurately assess current conditions. These lag times make nowcasting – or, the prediction of the present – an important practice. The problem is of greater importance in emerging markets, where lags in the release of data are often longer than in OECD countries. The availability of data that would allow decision-makers to observe trends as they unfold may improve the quality of economic assessments and, in turn, the decisions they inform.

In 2009, Google began the public release of its users' search queries through a publicly accessible interface. The rapid expansion of the Internet into all aspects of modern life together with Google's dominance in the search engine market give the company a central role in the collection of market intelligence. The publication of user search queries offers researchers the tantalizing possibility to observe the interests of society in real time without carrying out costly surveys.

The company's Chief Economist, Hal Varian, published a research note signaling the data's potential uses shortly after the public release of the Google Trends portal. The research question of immediate interest in the literature was whether the population's Internet tendencies contained a signal about their subsequent behaviour. Choi & Varian (2009b) show that basic nowcasting models for unemployment claims and vehicle sales in the United States can be substantially improved by incorporating search results for related keywords. While their analysis identifies promising improvements in model precision, they stop short of testing the hypothesis that Google data can improve a model's ability to identify turning points in the data. Extending the scope of the concept, Della Penna & Huang (2009) use search results for retail goods to nowcast private consumption in the United States, and find that their index is a better predictor than the commonly used Michigan Consumer Sentiment Index and Conference Board Consumer Confidence Index. Schmidt & Vosen (2009) carry out a similar analysis using a principal component approach and also find that their index outperforms the survey-based indexes currently in use.

Suhoy (2009) conducts a growth cycle analysis for Israel, an emerging market country that has an unusually high level of Internet penetration (71% in 2006), and twice the income per capita of any South American country. She investigates whether Google query indexes could have predicted the 2008 downturn in realtime, and finds that many web query categories do have predictive ability for their corresponding real activity.

So far the literature establishing the link between Internet search patterns and observed consumer behavior has been limited to advanced economies, and we are unaware of attempts to test this link in an emerging market where penetration rates and wealth levels are much lower. It is not immediately clear that the Internet has become embedded into the consumer's purchasing decision in emerging economies. While Internet use is growing rapidly in many South American countries, penetration is only half of what is observed in the United States and Western Europe (World Bank, 2009). Furthermore, a larger share of private consumption is taken up by non-discretionary spending on basic items that do not require extensive research prior to purchasing. As such, there is reason to believe that the information contained in Internet search queries may represent too small a sample of the total population to provide a valuable signal about aggregate behavior.

With these considerations in mind, we limit our analysis to the Chilean automotive sector for three reasons. First, the purchase of an automobile is a major event for most households, and one that requires a great deal of prior research into competing products. Second, the industry is dominated by a relatively small group of dedicated brands that can be adequately represented by a small number of keywords. And finally, Chile has one of the highest rates of Internet usage in Latin America at 35% of the population. The purpose of our empirical investigation is to evaluate the claim that Internet behavior correlates with consumer purchases in an emerging market. We do not aim to identify the most accurate model for nowcasting car sales, but rather to evaluate to what extent the popularity of Google search queries provides useful information about important macroeconomic variables in an emerging market.

2 Data

The search data is publicly available from Google through their Insights interface, and is available at a weekly frequency.[1] The interface returns one series per keyword for a given geographical area. For most countries, Google has also constructed categories that aggregate related keywords, but this feature has not yet been developed for Spanish-speaking countries.

The raw data undergoes two transformations prior to public release. First, the data are normalized by the total number of search queries in the geographical region of interest. As such, any trends from growth in the total number of Internet users or from a change in the relative popularity of Google as a search engine are removed from the data. Second, the normalized data are rescaled to an index with a maximum value of 100. This means that magnitudes are not directly comparable across series as a measure of relative popularity.

For our study, we use individual series on search queries for nine of the most popular automobile manufacturers in Chile by volume of sales, which together make up over 65% of sales by volume.[2] We use year-over-year changes in variables throughout the analysis for two reasons: (i) We are unable to reject the hypothesis of a unit root in the level of car sales, and (ii) To avoid spurious correlations from seasonal effects that our short sample length does not allow us to reliably model. Suhoy (2009) suggests that Google popularity series for consumption-related keywords may display a downward trend as the Internet becomes increasingly used for social-networking, and uses first-differences to avoid stationarity problems. While we observe no such trend in the series used in our analysis, we acknowledge the possibility of such a problem in theory and encourage practitioners to check the data prior to modelling.

[1] Available at http://www.google.com/insights/search/. Note that daily data is available for short sample lengths, but the sampling procedure employed by Google makes this data noisier than the weekly data.

[2] Market share data are from the Chilean National Automobile Association (ANAC). The keywords used are: Chevrolet, Hyundai, Nissan, Kia, Toyota, Suzuki, Ford, Mitsubishi, Mazda.

2.1 Addressing sampling noise

One problem with the data is that Google employs a sampling procedure that introduces noise into the series. The sampling appears to take place daily, such that requesting an identical query on different days returns slightly different series. If we assume that the underlying sampling procedure is unbiased, then downloading the series for each keyword on a number of occasions should allow us to identify the underlying signal in the data.

We downloaded each series on 17 occasions during the months of May and June 2010, and then computed the cross-sectional mean for each series at each time t. The resulting time series is then used as the historical series for each keyword. The upper section of figure 1 plots the 20th and 80th percentiles of the cross-sectional distributions, as well as the mean series for Chevrolet over the sample period. In the lower part of the figure, we plot the signal-to-noise ratio, or the ratio of the cross-sectional standard deviation to the mean. A higher value indicates that the data are providing a clearer underlying signal, and according to the Rose criterion used in the field of image processing, a value greater than 5 should be considered "high".

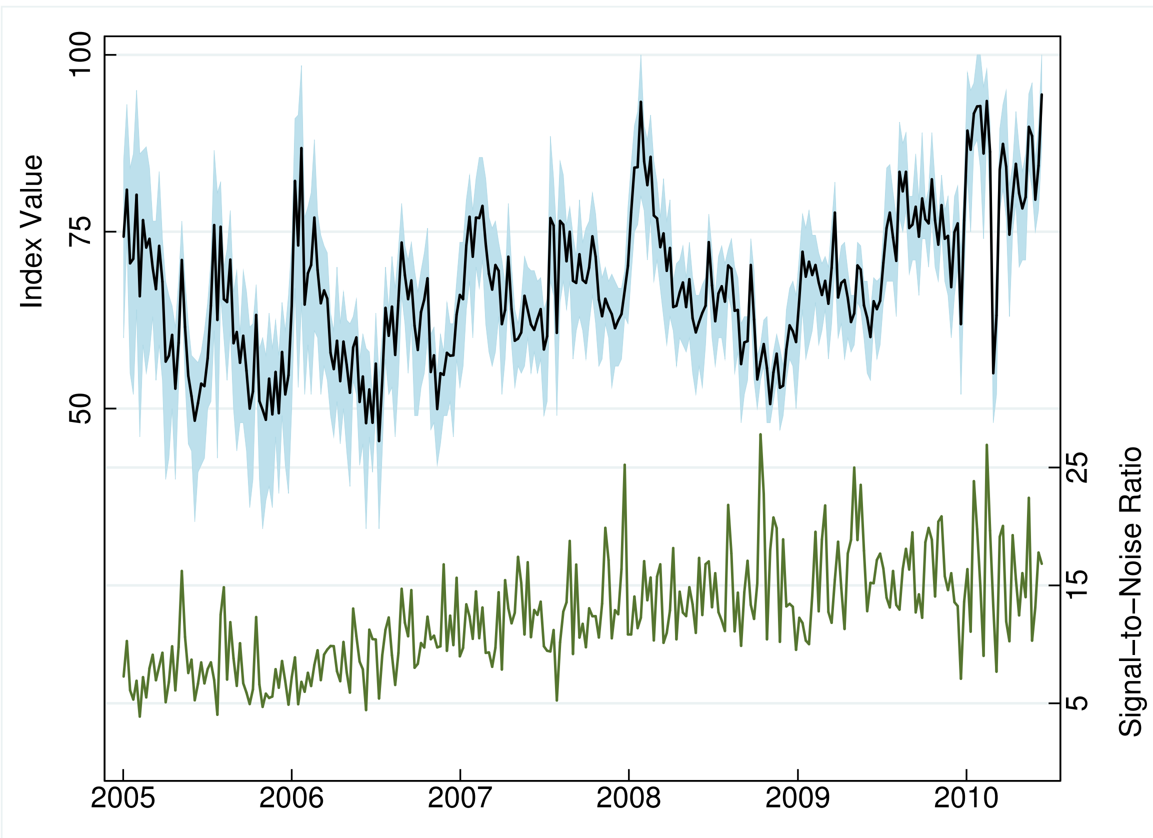

Figure 1: Sample noise: Series returned for keyword "Chevrolet"

2.2 Other data

Time-series data on the volume of car sales are from the national statistics agency, the *Instituto Nacional de Estadisticas de Chile*, and include sales of both new and used vehicles. After differencing the data, the available sample period used in our analysis is January 2006 to May 2010.

It should be noted that automobile sales are a volatile series. The coefficient of variation of the growth rate in sales, 1.96, is more than twice that of the growth rate in the industrial activity index (IMACEC). Figure 2 plots the series for the growth rate of sales and IMACEC on the same scale.

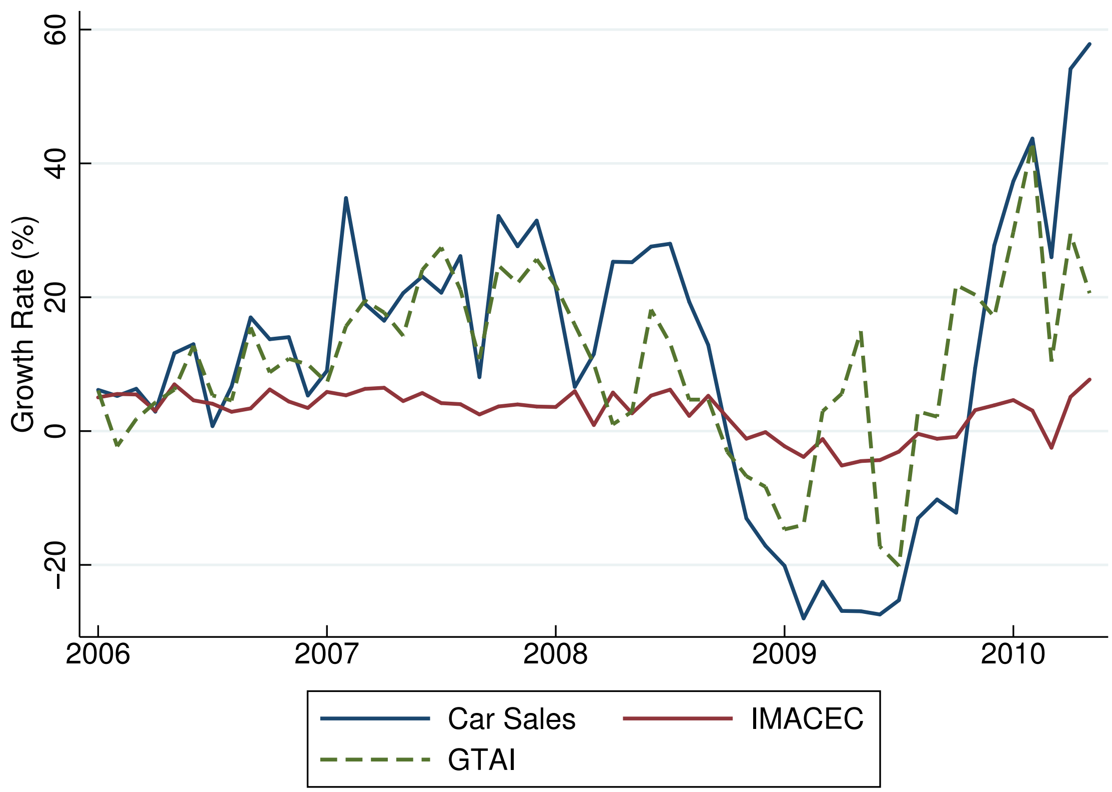

Figure 2: Data series employed

3 Methodology

In the absence of search categories for Chile, we are left with the task of aggregating individual keyword series into an index of automobile-related queries. Rather than using a method based on another source of data, such as the market share of each brand, we fit the series to a linear model and allow the weights to take on any value on the real line. Let $\mathbf{W}$ be a matrix of the year-over-year percentage change in the popularity of related keywords. We re-estimate the weights $\hat{\beta}$ in each period t by fitting the following model using the observations up to $t-1$:[3]

$$y_t = \alpha + \beta W_t + \epsilon_t$$

where y_t is the year-on-year percentage change in car sales and ϵ_t is a white noise error term. The Google Trends Automobile Index (GTAI) for period t is then computed as the fitted values from the regression, excluding the constant:

$$GTAI_t = E[\hat{\beta}|\Omega_{t-1}] \cdot W_t,$$

where Ω_t is the information set available at time t. Due to the transformations Google applies to the data before public release, our index is not a direct measure of search volume. Rather, it is the linear combination of the columns of $\mathbf{W}$ that best explains actual sales, and is thus more appropriately interpreted as the year-over-year percentage change in interest to purchase an automobile among Chilean Internet users.

We begin our empirical exercise by fitting linear autoregressive specifications of order $p = 1, ..., 6$ for car sales to determine the data's time-series structure. The parsimonious AR(1) specification provides a strong fit, which is reduced by adding higher-order autoregressive terms. An examination of model residuals reveals that no statistically significant autocorrelations remain, so we consider the AR(1) model an appropriate benchmark specification.

To ensure that the Google data are not simply providing information already available from other sources, in our second benchmark we augment the AR(1) specification by including the lagged values of the IMACEC, an index of aggregate economic activity in Chile that is commonly used due in forecast models due to its monthly frequency.

In the interest of encountering the strongest possible time-series characterization, while acknowledging the risk of over-fitting the data, we estimate a family of ARMA(p,q) models of order up to $p = 3$ and $q = 6$. Table 1 displays the results for competing specifications. The ARMA(2,2) is the preferred model on the basis of both the Akaike (AIC) and Bayesian Information Criterion (BIC), and is used as our final benchmark specification.

The three benchmark models are:

$$y_t = \alpha_{1a} + \rho_{1a} y_{t-1} + v_t \tag{1a}$$

$$y_t = \alpha_{2a} + \rho_{2a} y_{t-1} + \delta_{2a} z_{t-1} + \psi_t \tag{2a}$$

$$y_t = \alpha_{3a} + \sum_{p=1}^{2} \rho_{3a,p} y_{t-p} + \epsilon_t + \sum_{q=1}^{2} \theta_{3a,q} \epsilon_{t-q}, \tag{3a}$$

where z_t corresponds to the growth rate of the IMACEC index. We then introduce a family of augmented models that incorporate contemporaneous values of the GTAI variable:

[3]Note that the $GTAI$ for January 2006 to December 2007 is computed using all the data over the period.

Table 1: In-sample ARMA(p,q) Benchmarks

p	q	AIC	BIC	RMSE
1	1	406.9	414.8	10.3
1	2	408.8	418.7	10.3
1	3	410.0	421.9	10.2
1	4	410.4	424.2	10.1
1	5	404.6	420.4	9.3
1	6	412.1	427.8	10.0
2	1	408.9	418.8	10.3
2	2	397.5	407.3	9.1
2	3	398.6	410.4	9.1
2	4	400.5	414.3	9.0
2	5	406.6	424.3	9.3
2	6	403.8	421.5	9.0
3	1	401.4	413.2	9.3
3	2	398.8	410.6	9.1
3	3	400.5	414.3	9.0
3	4	402.1	417.9	9.0
3	5	406.8	426.6	9.1
3	6	398.9	418.6	8.4

$$y_t = \alpha_{1b} + \rho_{1b} y_{t-1} + \gamma_{1b} GTAI_t + \nu_t \tag{1b}$$

$$y_t = \alpha_{2b} + \rho_{2b} y_{t-1} + \delta_{2b} z_{t-1} + \gamma_{2b} GTAI_t + \varphi_t \tag{2b}$$

$$y_t = \alpha_{3b} + \sum_{p=1}^{2} \rho_{3b,p} y_{t-p} + \gamma_{3b} GTAI_t + \varepsilon_t + \sum_{q=1}^{2} \theta_{3b,q} \varepsilon_{t-q}. \tag{3b}$$

where υ_t, ψ_t, ϵ_t, ν_t, φ_t, and ε_t are white noise error terms.

We are interested in testing whether $GTAI$ contains relevant information beyond that which is contained in previous values of the dependent variable and other macroeconomic variables. This conceptual framework was proposed by Granger for determining the direction of causality between two variables.

There is a lively ongoing discussion in the field of economic forecasting over how best to test such a hypothesis. Ashley et. al (1980) argue that an out-of-sample prediction approach is more in the spirit of Granger's conceptual framework, since it only uses information available at the time the forecast is generated. Chen (2005) uses a Monte Carlo experiment to test the relative power of the two methods, and finds that out-of-sample tests do indeed offer higher power when the dependent variable exhibits a structural break. When no structural break is present, however, the in-sample test provides a higher power for testing the null hypothesis.

The short length of our sample ($T = 53$) makes it difficult to perform rigorous tests of parameter stability, but a visual inspection of parameter estimates $\hat{\gamma}$ and $\hat{\rho}$ from a recursive out-of-sample estimation of model (1b), plotted in figure 4 reveals substantial movements. Since our aim is limited to establishing the relevance of the $GTAI$ and not to identifying the best estimation procedure, we test the null hypothesis of no Granger causality using nowcasting results from an in-sample estimation and two out-of-sample estimation schemes to verify the robustness of our findings.

4 Results

4.1 In-sample Estimation

We begin our analysis by comparing the augmented models to their relevant benchmarks on the basis of in-sample predictive accuracy. To do so, we estimate model parameters using the complete sample and use the results to compute nowcast residuals. Estimation results are displayed in table 2. Model point accuracy is compared on the basis of the adjusted-R^2 and Root Mean Squared Error (RMSE), and skill at identifying the direction-of-change is compared using a sign test.

Since the benchmark models are nested within the augmented models, testing whether $GTAI_t$ Granger-causes y_t corresponds to a simple t-test of the null hypothesis that $\gamma = 0$ or, equivalently in bivariate models, whether $MSE_a \leq MSE_b$. As is reported in Table 2, the null is rejected for each of the augmented models. Furthermore, the introduction of Google information improves the fit over each reduced-form specification, reducing the RMSE between 2% and 13%. Since it generates the best fit and its residuals appear homoskedastic, we focus the remaining discussion on models (3a) and (3b), while pointing out that the results are broadly consistent across the three benchmarks.

Besides testing Granger causality on a point-accuracy basis, as we have done so far, it is of interest to test the related but distinct hypothesis that the $GTAI$ can identify turning points in the data. To test this claim, we conduct an event-forecast evaluation using the sign test methodology employed by Trefler (1995) and related test statistic proposed by Henriksson and Merton (1981). We begin by computing the share of periods in which each model accurately predicts the sign of the first and second differences in sales. We find that both models (3a) and (3b) correctly predict a rise or fall in car sales in over 90% of periods, but that model (3a) correctly identifies the direction of the change in the growth rate in only 50% of months. The introduction of Google Trends data in model (3b) increases the success rate to 67%. We verify whether the model's success rate is statistically significant using the market timing test proposed by Henriksson & Merton (1981). While model (3a) generates a test statistic below 1, indicating poor timing skill, we are able to reject the null hypothesis of no timing skills for model (3b) at the 95% confidence level.

These results are consistent across each of the benchmark models we have mentioned above. The substantial improvements in in-sample nowcast accuracy over the benchmark models suggests that the $GTAI$ indeed contains information beyond that which is contained in previous values of sales and macroeconomic variables.

4.2 The optimal window of search information

The importance of our findings is emphasized by comparing data release times. Data on actual car sales are released monthly by the National Statistics Agency during the last five days of the following month. Data from Google are available weekly with no delay in their release, such that complete Google data for a given month are available a full three weeks before the sales figure.

So far, we have included Google data available at time t and have compared models that use data in the information set available on the last day of each month to nowcast sales for that month. However, it is not necessarily the case that Internet users search for automobile information and then proceed with their purchase in the very same week. We investigate the possibility of a lag between searches and purchases that would suggest that Google Trends could be considered a slightly leading indicator.

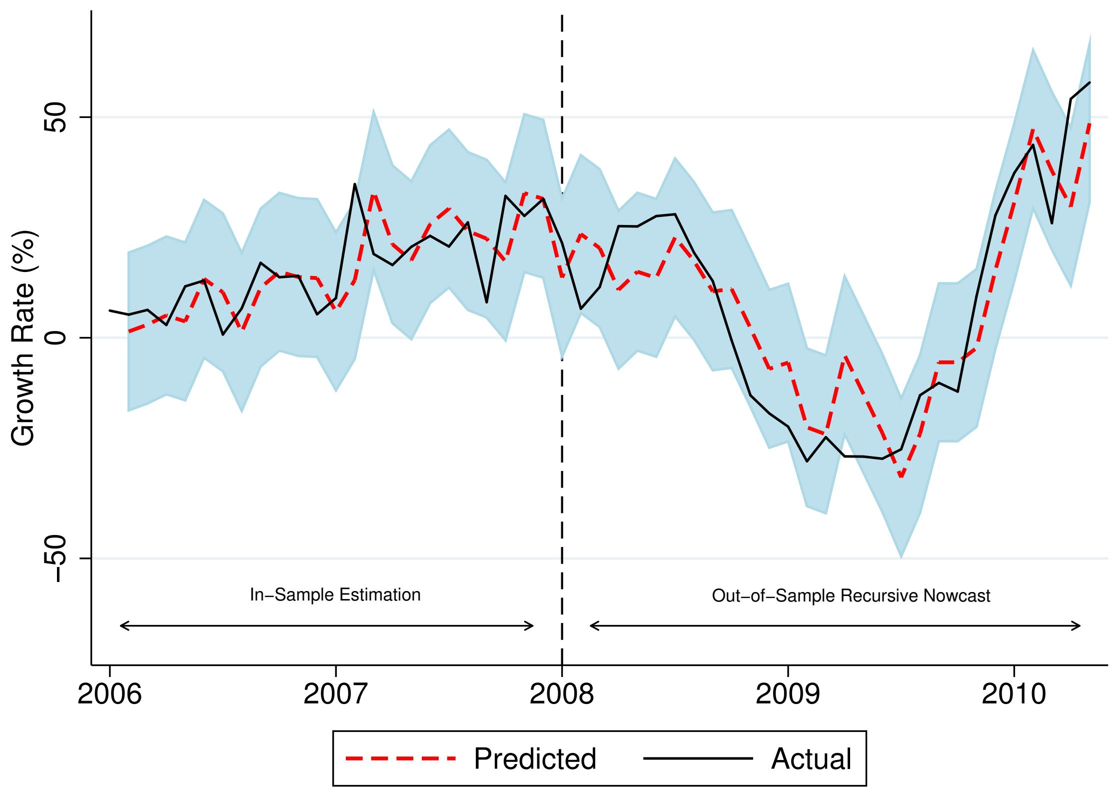

Figure 3: Nowcast for automobile sales using model (1b)

We re-estimate model (3b) having shifted a one-month window of search query observations backwards by one, two, and three weeks. Interestingly, the RMSE decreases and the adjusted-R^2 increases for lags of up to two weeks. Increasing the lag length further to three or four weeks reduces model fit, and lags of one month or more have no explanatory power. According to this underlying relationship between searches and purchases, a policymaker already has the best information for nowcasting automobile sales at her disposal one week prior to the end of the month, providing further gains in data release times. In fact, the inclusion of Google Trends information already improves nowcasts a full three weeks prior to the close of the month of interest. Using the optimal window of observations, the model correctly identifies turning points in the growth rate in 73% of periods.

Table 2: In-sample estimation results

Dependent Variable: Car sales (y_t)

	Benchmark Models			**Specifications Including GTAI**					
		Model			Model			Model	
Ind. Variable	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)		(3b)	
GTAI									
Contemporaneous				*0.54	*0.54	*0.54			
				(0.14)	(0.14)	(0.16)			
Shift 1 week							*0.48		
							(0.22)		
Shift 2 week								*0.42	
								(0.16)	
Shift 3 week									*0.42
									(0.16)
Car sales (y_t)									
$t-1$	*.91	*0.97		*0.70	*0.71				
	(.06)	(0.15)		(0.10)	(0.15)				
$ARMA(2,2)$			✓			✓	✓	✓	✓
IMACEC (z_t)									
$t-1$		-0.42			-0.04				
		(0.93)			(0.80)				
N	52	52	51	52	52	51	51	51	51
R^2	0.75	0.76	0.80	0.82	0.82	0.81	0.84	0.83	0.78
Adj. R^2	0.75	0.75	0.79	0.81	0.81	0.79	0.82	0.81	0.76
RMSE	10.5	10.6	9.14	9.12	9.22	8.97	8.22	8.46	9.60

Notes:
(i) All variables are 12-month percentage changes.
(ii) White heteroskedasticity-robust standard errors are reported below coefficient values.
(iii) All regressions include a constant term.
(iv) * Significant at the 0.01 level.

4.3 Out-of-Sample Nowcasts

The analysis presented thus far has demonstrated that the introduction of Google information outperforms univariate specifications in *ex-post* in-sample estimations. We now investigate the claim that Google data can also improve nowcast accuracy in out-of-sample estimations. We undertake the exercise using two estimation schemes common in the literature, each of which has advantages and shortcomings. Results are reported in table 3.

Under a recursive scheme, we begin by estimating the GTAI and model parameters over the first R periods, which corresponds to the interval of January 2006 to December 2007. These estimates are then used to formulate the first nowcast for period $R+1$. We then re-estimate the models in each period $t \in \{R+1, ..., T\}$ by extending the sample forward period by period, and computing the nowcast error for contemporaneous car sales in period $N+1$. This method has the virtue of using all the information available at time t, such that parameter estimates are expected to converge to the in-sample estimates as $N \rightarrow T$.

Under the recursive scheme, each of the benchmark models exhibits lower Mean-Squared Prediction Errors (MSPE) when the Google data is included. In the spirit of establishing Granger causality, we would like to test whether this improvement in forecast error is due to chance. Testing the relative performance of out-of-sample prediction models is a lively topic in the econometrics literature.

Since the benchmark models are nested within the augmented models, model residuals are identical (and thus, are not independent) under the null hypothesis of equal forecast performance. As a result, the test statistic proposed by Diebold and Mariano (1995) will not be normally distributed under the null, nor will the modified test statistic presented in Harvey et. al (1997) follow a Students-t distribution in small samples.

Testing the hypothesis of Granger causality is further complicated due to serial correlation in the forecast residuals. None of the models considered in this paper are able to adequately capture the serial correlation in the error terms following the start of the 2008 financial crisis. The ENC-NEW test statistic proposed by Clark and McCracken (2001) follows a known distribution in small samples, but this statistic may not follow the reported distribution when forecast errors display serial correlation.

Finally, we compute Clark & West's MSPE-adjusted test statistic, which is robust to a broad range of estimation methods and serial correlation in the error terms. We estimate the standard error of the test statistic using the Newey-West heteroskedasticity and autocorrelation consistent covariance matrix, and are able to reject the null hypothesis at the 90% confidence level.[4] The relative performance of each family of models is not maintained in the out-of-sample estimations. We find that the AR(1) specification in model (1a) produces more accurate out-of-sample nowcasts than does model (3a), so we focus our tests on the former. If the researcher were interested in identifying the best model to use, this inconsistency may be of concern. For our purposes, however, it suffices that the Google Trends-augmented models continue to outperform their respective reduced-form benchmarks.

As was alluded to in section 3, there is reason to believe that parameter estimates are

[4]See Clark & West (2006) for details on how to compute the test statistic.

Table 3: Out-of-Sample Causality Tests

Test	Models	Null hypothesis	Test Stat.	P-value
Recursive scheme:				
ENC-T	1a vs 1b	$MSE_a = MSE_b$	3.42	0.00
ENC-T	2a vs 2b	$MSE_a = MSE_b$	3.77	0.00
ENC-T	3a vs 3b	$MSE_a = MSE_b$	2.63	0.00
ENC-NEW	1a vs 1b	$MSE_a = MSE_b$	15.91	0.00
ENC-NEW	2a vs 2b	$MSE_a = MSE_b$	15.62	0.00
ENC-NEW	3a vs 3b	$MSE_a = MSE_b$	10.26	0.00
Clark & West (2007)	1a vs 1b	$MSPE_a = MSPE_b$	1.44	(0.05,0.10]
Clark & West (2007)	2a vs 2b	$MSPE_a = MSPE_b$	1.59	(0.05,0.10]
Clark & West (2007)	3a vs 3b	$MSPE_a = MSPE_b$	1.20	>0.10
Giacomini & White (2006)	1a vs 1b	$E[L_{a,t+1} - L_{b,t+1}\|t]$	7.3	0.03
Giacomini & White (2006)	2a vs 2b	$E[L_{a,t+1} - L_{b,t+1}\|t]$	7.08	0.03
Giacomini & White (2006)	3a vs 3b	$E[L_{a,t+1} - L_{b,t+1}\|t]$	2.54	0.28
Rolling window scheme:				
ENC-T	1a vs 1b	$MSE_a = MSE_b$	3.33	0.00
ENC-T	2a vs 2b	$MSE_a = MSE_b$	3.71	0.00
ENC-T	3a vs 3b	$MSE_a = MSE_b$	2.44	0.00
ENC-NEW	1a vs 1b	$MSE_a = MSE_b$	14.75	0.00
ENC-NEW	2a vs 2b	$MSE_a = MSE_b$	16.65	0.00
ENC-NEW	3a vs 3b	$MSE_a = MSE_b$	8.48	0.00
Clark & West (2007)	1a vs 1b	$MSPE_a = MSPE_b$	1.40	$(0.05, 0.10]$
Clark & West (2007)	2a vs 2b	$MSPE_a = MSPE_b$	1.60	$(0.05, 0.10]$
Clark & West (2007)	3a vs 3b	$MSPE_a = MSPE_b$	1.08	>0.10
Giacomini & White (2006)	1a vs 1b	$E[L_{a,t+1} - L_{b,t+1}\|t]$	7.37	0.03
Giacomini & White (2006)	2a vs 2b	$E[L_{a,t+1} - L_{b,t+1}\|t]$	6.34	0.04
Giacomini & White (2006)	3a vs 3b	$E[L_{a,t+1} - L_{b,t+1}\|t]$	2.03	0.36

Under the rolling scheme we estimate the models using a 24-month rolling window.

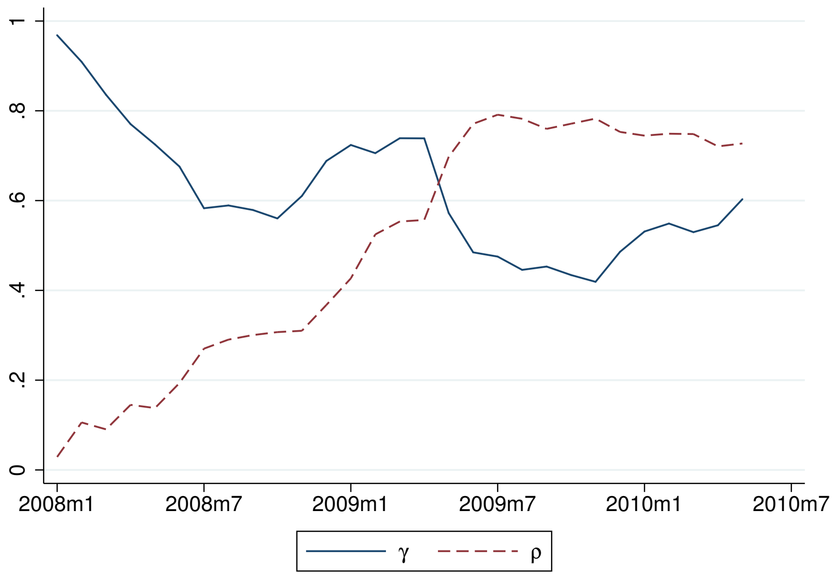

Figure 4: Estimated parameters from recursive out-of-sample nowcasts using model (1b)

unstable over the out-of-sample prediction period. A plot of estimated parameters from the recursive nowcast of model (1b) is shown in figure 4, where we observe a large change following the first few months of the global financial crisis, beginning approximately in September 2008, and possible convergence to new values approximately one year later. If this interpretation is correct, Chen (2006) and Giacomini & White (2007) point out that forecast accuracy and test power can be improved by using a rolling window estimation scheme, in which the model is re-estimated in each period $t \in (R+1, ..., T)$ using a sample of a fixed size.

We conduct a 24-month rolling window estimation to test the robustness of our findings. We again test the null hypothesis that model (1a) and model (1b) have equal population MSPE using Clark & West's MSPE-adjusted test, and reject the null hypothesis at the 90% confidence level. We also test the null hypothesis that the estimated MSE from models (1a) and (1b) are equal using Giacomini & White's Conditional Predictive Ability (CPA) test. Using a quadratic loss function, we are able to reject the null hypothesis at the 90% confidence level.

The evidence presented in this section confirms that the inclusion of information on Google search queries improves both the in- and out-of-sample accuracy of models for automobile sales.

5 Conclusion

We have examined whether Google search results provide relevant information about sales of automobiles in an emerging market. We constructed a Google Trends Automotive Index

(GTAI) using search queries for Chile, and included this index as a regressor in a family of simple nowcasting models. Our results show that models incorporating Google search results outperform competing benchmark specifications in both in- and out-of-sample nowcasting exercises. The Google data have a number of characteristics that should make them particularly attractive to decision-makers in emerging markets: (i) They are derived directly from micro user data; (ii) They contain information on a large proportion of Internet users, which is a far more extensive sample than is commonly employed by surveying agencies; and (iii) They are released at high frequency and at regular intervals. Our finding that the accuracy of nowcasting models for automobile sales in an emerging market can be improved using contemporaneous search patterns suggests that Google data is a promising source of information for nowcasting components of aggregate demand in short-run models, an exercise which we leave to future research.

References

[1] Ashley, R., C. Granger, & R. Schmalensee (1980) "Advertising and Aggregate Consumption: An Analysis of Causality," *Econometrica* 48, 1149-1167.

[2] Askitas, N. & K. F. Zimmermann (2009) "Google Econometrics and Unemployment Forecasting," *Applied Economics Quarterly* 55(2), 107-120.

[3] Chen, S. (2005) "A Note on In-Sample and Out-of-Sample Tests for Granger Causality," *Journal of Forecasting* 24, 453–464.

[4] Choi, H. & H. Varian (2009a) "Predicting the Present with Google Trends," Unpublished paper, Technical Report, Economics Research Group, Google.

[5] Choi, H. & H. Varian (2009b) "Predicting Initial Claims for Unemployment Benefits," Technical Report, Economics Research Group, Google.

[6] Clark, T. & M. W. McCracken (2001) "Tests of equal forecast accuracy and encompassing for nested models," *Journal of Econometrics* 105, 85-110.

[7] Clark, T. & K. West (2006) "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics* 138, 291-311.

[8] Della Penna, N. & H. Huang (2009) "Constructing Consumer Sentiment Index for U.S. Using Google Searches," Working Papers 2009-26, University of Alberta, Department of Economics.

[9] Diebold, F. & R. Mariano (1995) "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13(3), 253-263.

[10] Giacomini, R. & H. White (2006) "Tests of Conditional Predictive Ability," *Econometrica* 74(6), 1545-1578.

[11] Henriksson, R. & R. Merton (1981) "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills," *Journal of Business* 54(4), 513-533.

[12] Schmidt, T. & S. Vosen (2009) "Forecasting Private Consumption: Survey-based Indicators vs. Google Trends," *Ruhr Economic Papers* 155.

[13] Suhoy, T. (2009) "Query Indices and a 2008 Downturn: Israeli Data," Bank of Isreal, Discussion Paper No. 2009-06.

[14] Trefler, D. (1995) "The Case of the Missing Trade and Other Mysteries," *American Economic Review*, 85(5), 1029-1046.

[15] World Bank (2009) "World Development Indicators 2009 (CD-ROM)," The World Bank Group.

Documentos de Trabajo Banco Central de Chile	**Working Papers Central Bank of Chile**
NÚMEROS ANTERIORES	PAST ISSUES

La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica: www.bcentral.cl/esp/estpub/estudios/dtbc. Existe la posibilidad de solicitar una copia impresa con un costo de $500 si es dentro de Chile y US$12 si es para fuera de Chile. Las solicitudes se pueden hacer por fax: (56-2) 6702231 o a través de correo electrónico: bcch@bcentral.cl.

Working Papers in PDF format can be downloaded free of charge from: www.bcentral.cl/eng/stdpub/studies/workingpaper. Printed versions can be ordered individually for US$12 per copy (for orders inside Chile the charge is Ch$500.) Orders can be placed by fax: (56-2) 6702231 or e-mail: bcch@bcentral.cl.

DTBC-587 Julio 2010
Inflation Targeting in Financially Stable Economies: Has it been Flexible Enough?
Mauricio Calani, Kevin Cowan y Pablo García S.

DTBC-586 Junio 2010
Heterodox Central Banking
Luis Felipe Céspedes, Roberto Chang y Javier García-Cicco

DTBC-585 Junio 2010
Recessions and Financial Disruptions in Emerging Markets: A Bird's Eye View
Stijn Claessens, M. Ayhan Kose y Marco E. Terrones

DTBC-584 Junio 2010
The role of labor markets in structural change
Miguel Ricaurte

DTBC-583 Junio 2010
Firm Dynamics And Real Exchange Rate Fluctuations: Does Trade Openness Matter? Evidence From Mexico's Manufacturing Sector
Miguel Fuentes y Pablo Ibarrarán

DTBC-582 Junio 2010
Revisiting Overborrowing and its Policy Implications
Gianluca Benigno, Huigang Chen, Christopher Otrok, Alessandro Rebucci y Eric R. Young

DTBC-581 Mayo 2010
Fiscal Deficits, Debt, and Monetary Policy in a Liquidity Trap
Michael Devereux

DTBC-580 Mayo 2010
Anchors Away: How Fiscal Policy Can Undermine "Good" Monetary Policy
Eric Leeper

DTBC-579 Mayo 2010
The Credit Channel and Monetary Transmission in Brazil and Chile: A Structural VAR Approach
Luis Catão y Adrian Pagan

DTBC-578 Mayo 2010
Borrowing Constraints and Credit Demand
Jaime Ruiz-Tagle y Francis Vella

DTBC-577 Mayo 2010
Proyección Agregada y Desagregada del PIB Chileno con Procedimientos Automatizados de Series de Tiempo
Carlos Medel y Marcela Urrutia

DTBC-576 Mayo 2010
Competition And Stability In Banking
Xavier Vives

DTBC-575 Mayo 2010
The Global Financial Crisis
Franklin Allen y Elena Carletti

DTBC-574 Mayo 2010
The Determinants of Household Debt Default
Rodrigo Alfaro, Natalia Gallardo y Roberto Stein

DTBC-573 Abril 2010
On the Quantitative Effects of Unconventional Monetary Policies
Javier García-Cicco